

Correct interpretation of trace normalized density matrices as ensembles

Paul M. Sheldon
(22 June 1996)

A density operator, $\rho = P_\alpha |\alpha\rangle\langle\alpha| + P_\beta |\beta\rangle\langle\beta|$, with P_α and P_β linearly independent normalized wave functions, must be traced normalized, so $P_\beta = 1 - P_\alpha$. However, unless $\langle\alpha|\beta\rangle = 0$, P_α and P_β cannot be interpreted as probabilities of finding $|\alpha\rangle$ and $|\beta\rangle$ respectively. We show that a density matrix comprised of two (P_α and P_β nonzero) non-orthogonal projectors have unique spectral decomposition into diagonal form with orthogonal projectors. Only then, according to axioms of Von Neumann and Fock, can we have probability interpretation of that density matrix, only then can the diagonal elements be interpreted as probabilities of an ensemble. Those probabilities on the diagonal are not P_α and P_β . Further, only in the case of orthogonal projectors can we have the degenerate situation in which multiple ensembles are permitted.

I. INTRODUCTION :

This paper is an attempt to elucidate the well known Schrodinger cat paradox with reference to its debate in Sir Roger Penrose and Stephen Hawking's recent joint book, "The Nature of Space and Time", by more fully presenting decoherences view that the wave function has merely an interpretation in probability rather than a definition in probability.

Even though, throughout the text, Penrose and Hawking publicly debate their views on the paradox, the casual reader could mistakenly conclude that just one page of Penroses algebra, page 70, effectively evaluates and dismisses the promising young program of decoherence.

Penrose uses both wave function and density matrix algebraic representation of the so called correlated state of Schrodinger's cat coupled to a memory state. Penrose claims decoherence doesn't explain the predicament of Schrodinger's cat because decoherence does not select from what he defines as two alternative ensembles.

Penrose postulates that these alternative ensembles are identical to weighted sums of wave function or ketbra products. But, as Zeh says (page 2 of his October 96 quant-ph entry), these are not ensembles at all :

This density matrix has a form *as though* it represented an ensemble of wave functions, ϕ_n^{local} , with respective probabilities, p_n . Equation (2) demonstrates, however, that it does not. This confusion between proper and "improper" mixtures has given rise to the most frequent misinterpretation of decoherence as leading to ensembles, and thus as deriving the collapse of the wave function as a stochastic process.

Decoherence never claimed to select between the things that Penrose calls ensembles.

Penroses "ensemble" 1 is [cat dead] \times [observed dead memory state] or [cat alive] \times [observed alive memory state] presumably orthogonal and without any overlap and "ensemble"

2 is two different superpositions of dead and alive (in products of observed and recorded state) that add up to the same thing in both state and density matrix representations.

Postulating this, doesn't lead to superselection.

Penrose, in the book, represents decoherence as an algebraic recognition about ways to arrange sums of products of factors. But, only a dynamical process will truly represent it. It is not, as in algebra, instantaneous, but instead happening over time (decoherence time). Decoherence only derives its power from the physical interaction between those factors, an interaction not addressed in Penroses algebra.

We next consider the physical meaning of the Schmidt polar form of orthogonal states, a purely formal device which predates any notions of quantum mechanics and certainly doesn't deal with the essential environmental interaction in the extended Von Neumann measurement model.

Penrose seems to imply that the orthogonality imposed by that form both is decoherences only claim to explanatory power and weakens the generality.

Regardless of measurement interaction, at each instant, one can find a sum of orthogonal wave function products called a Schmidt polar form. This, in itself, represents neither restriction (information about physical law or principle) nor loss of generality.

While the Schmidt polar form together with the orthogonality of the products will be required for superselection according to decoherence, that form alone isn't sufficient for the apparent gain of information by an experiment.

Rather, as Albrecht showed, for a sufficiently "large" measurement apparatus, interaction would, after a time, produce dynamically stable (to repeated measurement interaction) orthogonal states in a Schmidt polar form to counter the instability caused by nearly degenerate density matrix eigenvalues (probabilities).

Penrose sets up an extreme situation of precisely degenerate eigenvalues on p.70, the interpreted probabilities being mathematically prepared as exactly equal.

I claim it is wrong to imagine oneself able to so determine an initial wave function.

To see measurement as a collapse into a definite outcome with a certain probability requires that a repetition get that same outcome.

I first consider two systems in quantum mechanics. I let one system be an observers memory of an experiment and the other system be what is being observed. Quantum mechanics not only allows superposition, but the dynamics of the wave equation demands it. So we not only may consider the so called correlated state of these two systems (the superposition of the product pairs of a state from the first and a state from the second) but we must consider it.

In fact, according to Wheeler and Everett who addressed the crisis of interpretation produced by quantum cosmology in the absence of external observer, etc.), we must deal with several copies of observers observing different universes.

The unextended Von Neumann model (discussed in the last two pages of his 1930 monograph) already explained the unitary dynamics of measurement.

To impose time asymmetry, Von Neumann starts with a special initial state, a single product. He then shows, the measurement interaction will make that single product “branch” into a superposition of products. But, which products, Penrose might ask?

The function subspace of the observed system can be broken up into a special orthogonal decomposition that corresponds to the eigenvalues of a Hermitian operator, an observable.

A measurement apparatus, though its design was later recognized to be curtailed by the environment, must measure some such eigenvalue. Associated with the apparatus, according to Von Neumann, is an interaction Hamiltonian, comprised of two factors, one from the observed system, and one from the pointer. The former, with the eigenvalue we want to observe, is multiplied by the momentum of the pointer. The interaction Hamiltonian, suitably placed in an exponent, creates a time-evolution operator in the so-called interaction picture. For each eigenfunction building the observed system, this operator simplifies to an excursion operator on the pointer. And it distributes linearly over the system pointer wavefunction to produce what Zurek calls only the first stage of measurement, the correlated state. This correlated state is a superposition of products of an eigenfunction of the system factor and eigenfunction of the pointer position.

This “first-stage” correlated state can be rewritten algebraically to reflect alternative superpositions. That is why, Zurek says, it is incomplete and only the first stage.

The next stage occurs when you “record” the response of the pointer into your history and it remains consistently recorded, something that essentially involves the huge-dimensioned environment outside the pointer and measured system. This is the second stage of measurement, decoherence.

You are looking at the state multiplied by that pointer value in the correlated state sum, because (from an “outside” multiverse point of view, a point of view that paradoxically doesn’t exist but we talk about anyway) you are there in one of the products to view the other factor. Environmental decoherence doesn’t allow a viewpoint where you see correlated states for “macroscopic” things which are very entangled with the environment. Instead it selects a Schmidt polar form where all observers that see each other get on the same product and agree on the recorded history in that product. Communicate with any of Wigner’s friends and you will agree on a measurement result.

Relative to one particular you something happened, a collapse. The features of collapse are not only a moment by moment probability but also a permanence of record of what happened.

Relative to the multiverse nothing really happened but everything just is “there”, the “old” you splitting relative to some clock pointer position onto many products.

So, relative to you, before you record the response of the pointer, you must have something like a probability of what you would record later. What allows the interpretation of this probability-like thing as a probability?

Trace out the environmental and system degrees of freedom from the full density matrix of the pure state of the universe and you will get a reduced density matrix of the pointer. This reduced density matrix is trace normalized and can be orthogonalized to have its positive definite eigenvalues interpreted as probabilities of pointer positions. These probabilities turn out exactly the same as those associated with system states. How can this happen?

Expectations in the above so called improper mixtures have identical formal appearance to those in proper mixtures and the identical formal appearance of expectations lead us to speak of probabilities in “improper” circumstances.

The reduced density matrix can be used to define expectation values of operators, exactly as one built from real probabilities can. Not only that, consider finding the expectation of a special class of operators, the orthogonal projection ketbras.

These special expectations can define the probability of distinct eigenvalues of a Hermitian operator. Consider the spectral decomposition of that operator into projections times eigenvalues. The expectation value operator is linear, so that it can not only be distributed, but also pulled all the way out onto the individual projections. Now one sees that this expectation value of the operator conforms to the ordinary probability interpretation of expectation value of eigenvalues times probabilities of eigenstate.

In addition, for a probability interpretation of the outcome of a measurement, a repeated measurement must yield a certitude.

The eigenvalues of the reduced density matrix of the pointer are what are interpreted as probabilities of the observable’s operator if the environments measurement-like interaction doesn’t change the observables eigenvalue, i.e. if its measurement-like interaction is a quantum non-demolition one relative to the observables operator.

Whether or not the eigenvalues of the Hermitian density matrix of the pointer are degenerate, we can always construct an orthogonal set of eigenvectors.

What about other orthogonalities? By definition of Von Neumann measurement interaction with Hermitian operator, the dynamics of branching resolves orthogonal system states or subspaces. These distribute themselves among the terms of the correlated state produced by the excursion operator. Now consider other factors in each of these terms, namely the state of the observer, bath (environment), and discrete pointer position. These factors are only after decoherence time almost orthogonal. Albrecht provides an example showing that when one has a very large-dimensional apparatus, one can afford to have pointer probabilities closer to each other and still accept that the states produced by the excursion operator are close to providing the required orthogonal states of the Schmidt polar form.

Why then is this dimension enormous and how does this enormity allow almost orthogonality rather than the ambiguities of Zurek’s so-called first stage of measurement?

The physical pointers degrees of freedom are really combined with the environments degrees of freedom, e.g. photons that scatter off it. The photons are producing correlated states with the physical pointer and thus are themselves engaged in a measurement interaction along with the observer

correlating to the photons.

In dealing with a random measurement interaction , Albrecht points out that one large dimensioned unit vector randomly kicked by measurement around the unit sphere will have probably negligible projection on any one arbitrary axis and hence is probably almost orthogonal to that arbitrary axis. So, augmented by such a random environmental interaction, and after a so called decoherence time, which allows the dimension to get large enough, the observer states (pointer and environment) are almost orthogonal in a Schmidt polar decomposition of product pairs of orthogonal states.

The above example with complete randomness in the environment interaction makes plausible this almost orthogonality in a simple extreme way. When that interaction is not completely random, Zeh points out that one has examples of superselection. This means that certain things we propose to measure with an interaction might be erased by the environment.

Take, for example, macroscopic parity nonconservation. Due to the enormous interaction size with an environment of chirality-measuring photon polarization states, which erase parity, we had best not set our pointer to measure parity! On the other hand, measuring chirality repeatedly yields idempotency (along a branch) and thus can be interpreted as having recorded a certitude where there had been an ensemble. One then has a consistent history! Decoherence thus explains the puzzling fact that sugars are found in definite chiralities.

There are many other examples of such superselections which are well explained by decoherence.

Using only decoherences “traceout” feature to reduce the density matrix will fail to superselect between those ensembles which are supposedly more generally defined by Penrose et al. By using merely this feature as generating the probability, the casual reader of Penroses position might be misled to believe all the physics can be encrypted in an arbitrary and more general reduced density matrix.

I will illustrate, at first, with Penroses simple two-state system and memory, that his ensembles arent really more general. Penrose has implied this greater generality with the words “not necessarily orthogonal” in his final retort to Hawking on p.134. A referee of one of my recent papers had refuted this when he stipulated orthogonality as an obviously required condition for “interpretation as” (as distinct from “being”) probabilities.

Penrose cites a paper on ostensibly more general rho-ensembles, which ignores physical interaction dynamics, in contrast to proponents of decoherence, including the above mentioned referee.

By using Penroses postulate of what an ensemble is before decoherence-time rather than talking about something else that is interpreted as an ensemble after decoherence-time, we and not nature must choose the eigenensemble.

I will illustrate here that we can WOLOG use “eigenensembles”.

In fact, I also propose, in thesis, to study what else might be necessary “on the way to sufficiency” for probability interpretation. I use this phrase because probability interpretation (measurement resolvability of a superselected ensemble) is something only approached approximately in a so-called (e folding) “decoherence time”.

This physics dealing with time is not encrypted in the reduced density matrix! The reduced density matrix is not an ensemble. The Von Neumann operational definition of measurement and formally an infinite amount of time allows us to interpret what is not an ensemble as an ensemble and to see a special eigenensemble as superselected.

That is what the decoherence research has been illustrating in recent years and what seems not to be illustrated, but rather misrepresented, by Penroses algebra on p.70.

A finite time of apparatus-environment interaction only fails to superselect between eigenensembles that are too close to each other. “Closeness” of ensembles involves two factors, the smallness of the difference of eigenvalues of the reduced density matrix (eigenensemble probabilities) and also the size of the measuring apparatus-environment (entanglement dimensionality that increases in time with each branching due to another measurement).

“Buying a theory” is motivated by both cost (complication) and payoff (use). Buying decoherence is no different.

When a person “buys” a theory or way of interpreting things, he pays what I call Occams cost of complication. Occams razor would have us choose the interpretation of greatest simplicity.

Occams razor motivates our choice of a theory by its cost alone. The cost of decoherence is the introduction of unobservable universes (using only observed and tested quantum dynamics) while the cost of collapse is more laws (to cut out extra universes). Occams razor alone cant choose between them.

The advocate of decoherence, however, gets a payoff. He may design his instrument to either resolve or blur close by eigenensembles. He can blur by looking at things before decoherence time or resolve by looking after. The collapse conjecture of Penrose doesnt seem to have yet predicted any such time with hard numbers like Zurek has.

If the advocate of decoherence still wants to see these alternate ensembles after a time of entanglement (e.g. quantum computers after the required number of machine cycles), he designs that apparatus to remain in that time of small size by isolating it from the environment interactions, for example, by refrigeration. Then, the decoherence time limit will allow the decoders enough quantum computer machine cycles to solve the cryptographic factoring problem.

If the advocate of decoherence wants to rapidly see whether Schrodingers cat is alive or dead, he lets the apparatus get large in the limited time allowed by the experimental schedule.

To have similar design options, an advocate of collapse must yet find a collapse time where Zurek and others have already found a decoherence time. How easy will that be to do?

Consider some excerpts from Penroses The Emperors New Mind (on p.367, 368, 369 and 372) :

...

At some stage, this complex quantum linear superposition becomes a real probability-weighted collection of actual alternatives ... According to the viewpoint that I am proposing, that stage occurs as soon as the difference between the gravitational fields of the various al-

ternatives reaches the one-graviton level.

...

How ‘big’ is the one graviton level? ...more a question of mass and energy distribution. ...The characteristic quantum-gravitational scale of mass is what is known as the Planck mass ...These must be allowed to be what are called longitudinal gravitons - the virtual gravitons that compose a static gravitational field. Unfortunately, there are theoretical problems involved in defining such things

...

(Italics are mine)

I brief the above :

To determine a time of collapse, one must find a method of counting to a very obscure **one**. For, collapse will evidently happen when there is enough curvature difference between would-be superposed states, a difference of environmental curvature related vaguely to **one** (virtual and therefore possibly massive) exchange-force-graviton of about Planck-mass.

In contrast to predicted and therefore experimentally testable decoherence time in linear quantum mechanics, where a size is found by merely counting the dimensions of an enlarging apparatus-environment, evidently a collapse time will be for quite some time immune to verification.

On an interesting side note, H. D. Zeh pointed out to me a publication by Kiefer, a decoherence advocate. Kiefer treated therein real gravitons causing decoherence. Treating vacuum solution minisuperspaces (ones that only contain gravity coupled to nothing else), Kiefer attempted to show therein that real gravitons cause decoherence, superselect them as classical and render a classical, though so-called internal, time.

John Wheeler, an advocate of decoherence, pointed out in Batelle Rencontres the crisis in **interpretation** of quantum mechanics brought out by the new field of quantum cosmology, the absence of an external observer (and external time).

Yet, an advocate of collapse would sell us a complication of theory at what would be the final (somehow macroscopic) point in the correlation chain, consciousness (better than any currently understood computer). Here he would demand an extended quantum theory that gives actual ensemble states and not a single consciousness correlated state. He would demand **not an interpreting**, participating, branching consciousness, but one single **external** consciousness, perhaps an isolated quantum computer that could observe.

Since any one really correlated consciousness would see one reality or another, what difference is there to saying, by a sort of “relative-to-me” principle, that the other realities aren’t there?

The difference isn’t seen by Occam’s razor.

But an advocate of decoherence has the payoff that he can recognize that an isolated quantum computer not isolated from the observation process is a contradiction in terms. I maintain that the collapse advocate puts together too much in one system. He puts together the time irreversible recording of computational results with what Fredkin and others have shown to be a fundamentally time reversible quantum computational process.

So, rather than classical collapse and “relative-to-me” possibilities being the ultimate arbiter of what is, we might, in-

stead, through quantum principles design things as we please, the classical or nonclassical, computer reversible process or irreversible record, and perhaps even time or nontime.

A recent Texas Instruments meeting affirmed that the power to specify options like this goes beyond “mere” physics curiosity.

Seth Lloyd, a mechanical engineer at MIT, demonstrated (Science 17 Sept 1993, Vol.261, pp.1569-1571) that coherent Pi-pulses can both enter data into, and manipulate, quantum two state systems (so called q-bit registers).

In discussion following the meeting, we agreed in the usefulness for these computers of my proposed study of superselected states and the derivation and design of their classicality. He maintained that his coherent Pi-pulses, because they are most classical, are most suited as “probes” of (as distinct from measurements in) quantum mechanics. For clarity, let’s call such a probe a “q-observer” to distinguish it from an observer that records measurements. We see only the first stage of measurement! How is this so?

I had already followed the mathematics of a seminal paper, (Zurek, Physics of Time Asymmetry, “Preferred sets of states, predictability, classicality, and environment-induced decoherence” p.175-212) and found that coherent state classicality comes from having the least Von Neumann entropy production, the least decoherence. That is, we might say classicality almost happens when decoherence almost stops. Ironically, that also implies the least disturbance of what is left of the quantum mechanical.

Seth Lloyd and I both realized that this minimum Von Neumann entropy production of a superselected classical state could have a dual role to play.

We cited Caves work with density matrices and mutual information considerations and saw that, were a q-observer and the q-observed designed in isolated interaction starting from an initial pure state, their reduced density matrices would show identical entropies. Minimum Von Neumann entropy production by a classical state would guarantee the same for the state which it probes.

So, the dual roles are first for the state itself and second for a quantum mechanical thing with which it alone interacts.

While a most classical state, say a coherent Pi-pulse, is the least measured i.e. decohered by the environment, it also measures the least i.e. decoheres a quantum computer the least, thus disturbing the benefits of their quantum superposition i.e. parallelism the least.

Such classical (collapsed as much as possible) states have a criterion for payoff, a criterion based on a predictable time limit (for quantum computing within), namely the decoherence time. Decoherence fulfills the Popper requirement of a theory, providing predictions that can be falsified by experiment. If that decoherence time consistently proves to be a criterion of payoff, that would weigh heavily for decoherence theory rather than collapse theory which offers no such payoff.

So, I would study generalized coherent states to better understand what is really going on with classicality for more reliable quantum computer probes.

II. ANALYSIS :

Without loss of generality of projector decomposition, we let phase of $|\alpha\rangle$ and $|\beta\rangle$ such that $\langle\alpha|\beta\rangle$ is real and positive :

$$\langle\alpha|\beta\rangle = |\langle\alpha|\beta\rangle| = c \neq 0 \quad (2.1)$$

Now, because $|\alpha\rangle$ and $|\beta\rangle$ are linearly independent, each eigenvector may be expressed as a unique linear combination of $|\alpha\rangle$ and $|\beta\rangle$. Let us express normalized eigenvectors, $|\hat{e}_1\rangle$ and $|\hat{e}_2\rangle$, in terms of $|\alpha\rangle$ and $|\beta\rangle$:

$$\begin{aligned} |\hat{e}_1\rangle &= c_{1\alpha} |\alpha\rangle + c_{1\beta} |\beta\rangle \\ |\hat{e}_2\rangle &= c_{2\alpha} |\alpha\rangle + c_{2\beta} |\beta\rangle \end{aligned} \quad (2.2)$$

We will now investigate a proposition, that one of $|\alpha\rangle$ and $|\beta\rangle$ appears in both linear combinations for the eigenvectors. We suppose the contrary and will achieve a contradiction.

If $c_{1\beta} = c_{2\alpha} = 0$, then each of $|\alpha\rangle$ and $|\beta\rangle$ appears in only one linear combination. Then we may as well write :

$$\begin{aligned} |\hat{e}_1\rangle &= |\alpha\rangle \\ |\hat{e}_2\rangle &= |\beta\rangle \end{aligned} \quad (2.3)$$

Applying the density operator to $|\hat{e}_1\rangle$, we find :

$$\begin{aligned} &P_{\hat{e}_1} |\alpha\rangle \\ &= P_{\hat{e}_1} |\hat{e}_1\rangle \\ &= \rho |\hat{e}_1\rangle \\ &= (P_\alpha |\alpha\rangle \langle\alpha| + P_\beta |\beta\rangle \langle\beta|) |\alpha\rangle \quad (2.4) \\ &= P_\alpha |\alpha\rangle \langle\alpha| |\alpha\rangle + P_\beta |\beta\rangle \langle\beta| |\alpha\rangle \\ &= P_\alpha |\alpha\rangle + c P_\beta |\beta\rangle \\ &\Rightarrow P_\beta = 0 \end{aligned}$$

But, this contradicts what we had assumed.

Therefore, one of $\{|\hat{e}_1\rangle, |\hat{e}_2\rangle\}$ appears in both linear combinations for eigenvectors (one of $c_{1\beta}$ or $c_{2\alpha} \neq 0$) and WOLOG we relabel and make that be $|\alpha\rangle$. We can then simply, with a suitable choice of phase, write a system of unnormalized eigenvectors where the coefficient of $|\alpha\rangle$ is one and r_j is defined to make the coefficient of the other be $c r_j$. Our system of unnormalized eigenvectors becomes :

$$\begin{aligned} |e_j\rangle &= |\alpha\rangle + c r_j |\beta\rangle \\ \text{where } |\hat{e}_j\rangle &= \frac{|e_j\rangle}{\sqrt{\langle e_j | e_j \rangle}} \end{aligned} \quad (2.5)$$

Surpressing the index j, we tabulate some results for later use :

$$\begin{aligned} \langle\alpha|e\rangle &= \langle\alpha|\alpha\rangle + c r \langle\alpha|\beta\rangle = 1 + c^2 r \\ \langle\beta|e\rangle &= \langle\beta|\alpha\rangle + c r \langle\beta|\beta\rangle = c (1 + r) \end{aligned} \quad (2.6)$$

Imposing the condition that $|e\rangle$ is an eigenvector of ρ , we find that :

$$\begin{aligned} &P_e (\langle\alpha| + c r \langle\beta|) \\ &= P_e |e\rangle = \rho |e\rangle = (P_\alpha |\alpha\rangle \langle\alpha| + P_\beta |\beta\rangle \langle\beta|) |e\rangle \\ &= P_\alpha |\alpha\rangle \langle\alpha| |e\rangle + P_\beta |\beta\rangle \langle\beta| |e\rangle \\ &= P_\alpha |\alpha\rangle (1 + c^2 r) + P_\beta |\beta\rangle c (1 + r) \end{aligned} \quad (2.7)$$

By equating above coefficients of linearly independent $|\alpha\rangle$ and $|\beta\rangle$, we obtain a system :

$$\begin{aligned} P_e &= P_\alpha (1 + c^2 r) \\ P_e c r &= P_\beta c (1 + r) \end{aligned} \quad (2.8)$$

Now, r can't equal zero, because then the second equation would imply that P_β is zero for contradiction to our assumptions. Further, r must be real to render a real probability in this equation. Then, straight off we see from first equation that our eigenvalues, probabilities, can't ever be equal to P_α (unless, contrary to assumptions, $c=0$) and they further are distinct.

Thence, without loss of information, we multiply both sides of first equation by cr to gain a common left hand side :

$$\begin{aligned} P_e c r &= P_\alpha (1 + c^2 r) c r \\ P_e c r &= P_\beta c (1 + r) \end{aligned} \quad (2.9)$$

which we eliminate for one quadratic equation in r :

$$P_\alpha (1 + c^2 r) c r = P_\beta c (1 + r) \quad (2.10)$$

We rewrite the quadratic in standard form :

$$c^2 r^2 + \left(1 - \frac{P_\beta}{P_\alpha}\right) r - \frac{P_\beta}{P_\alpha} = 0 \quad (2.11)$$

and invent an abbreviated notation :

$$c^2 r^2 + (1 - Prat) r - Prat = 0 \quad (2.12)$$

Our roots are :

$$r_{\pm} = \frac{Prat - 1 \pm \sqrt{(1 - Prat)^2 + 4 c^2 Prat}}{2 c^2} \quad (2.13)$$

Verify, from above roots, expected orthogonality (Hermitian operator of distinct eigenvalues) :

$$\begin{aligned} &\langle e_2 | e_1 \rangle \\ &= (\langle\alpha| + c r_2 \langle\beta|) (\langle\alpha| + c r_1 \langle\beta|) \\ &= \left(\begin{array}{l} (\langle\alpha|\alpha\rangle + c r_2 \langle\alpha|\beta\rangle) \\ + c r_2 (\langle\beta|\alpha\rangle + c r_1 \langle\beta|\beta\rangle) \end{array} \right) \\ &= ((1 + c r_1 c) + c r_2 (c + c r_1 1)) \\ &= 1 + (r_1 + r_2) c^2 + r_1 r_2 c^2 \\ &= 1 + (Prat - 1) - Prat = 0 \end{aligned} \quad (2.14)$$

Suppose the degenerate case (for eigenvalues= 1/2), then, from the first equation of the system from the eigenvector condition, we have root degeneracy equating the radicand to zero :

$$(1 - Prat)^2 + 4 c^2 Prat = 0 \quad (2.15)$$

c was assumed positive WOLOG, so solving c we get one root :

$$c = \frac{1}{2} \sqrt{-\frac{1}{Prat} - Prat + 2} \quad (2.16)$$

Now, only were radicand above positive definite would c be positive definite. Differentiation shows one extremum ($Prat = -1$ extraneous as not trace normalized).

$$\begin{aligned} y &= -\frac{1}{Prat} - Prat + 2 \\ 0 &= \frac{\partial y}{\partial Prat} = \frac{1}{Prat^2} - 1 \\ Prat &= 1 \end{aligned} \quad (2.17)$$

then plugging in
 $y = 0$ and $c = 0$

We observe radicand = - for $Prat = 0$ or , so that extremum is a maximum wherein $c = 0$.

III. CONCLUSION :

Spectral decomposition gives probability interpretation of mixture of projection operators. Degeneracy of such ensemble solution is only afforded in special case where projectors are both orthogonal and precisely equal probability.

IV. SUGGESTIONS FOR FURTHER READING :

I have constructed herein a formal, hopefully reassuring, brief on degeneracy of density matrix versus ensemble superselection.

Such formal briefs, while surveyable and publishable, can't really get the flavor of the story, the context, the references, the bigger perspective.

So, I invite the reader to a perspective that not merely from density matrix clarifies the formal probability definition, but exemplifies, in measurement apparatus, probability's wisely designed implementation.

For example, I have since my paper come across a delightful paper by Albrecht [1] who helped me relate the terms "macroscopic measurement apparatus" to this degeneracy. He finds increasing a measuring instrument's "size" (Hilbert subspace dimensionality) could make superselection insensitive to a "degree" of degeneracy.

Albrecht points out another feature important to implement correctly. Measurement interaction should insure (QND) stability of the measured factors in the products that superpose in the Schmidt correlated states.

Consider this stability.

While Scully in an invited tutorial paper [2] connects different people's notions of the coherent (classical superselected) states, Zurek's pointer basis [4] speaks of this stability against environmental measurement decoherence of these same states [5].

Finally, we "need the story" recorded by this stability!

Although reconstruction formidable [3], the essential quantum reality of the observed system before an experiment does not have to be disturbed or held as specious by such a well designed apparatus, in fact, throughout (both before and after) the experiment, the dynamics of the quantum reality enables that experiment's design.

V. ACKNOWLEDGEMENTS :

I would like to acknowledge long correspondence with Ulrich Gerlach about his class notes needed to help me attempt to understand Hawking's papers, "Origin of Structure" and "Origin of Time Asymmetry", and thereby WKB time, a term coined by Dr. C. Kiefer. Dr. C. Kiefer long ago corresponded with me about decoherence and time. I am grateful for an even longer correspondence with H. D. Zeh on many other notions of decoherence and for his preprints and suggested readings. I am most happy to have met Seth Lloyd who encouraged me at Texas Instruments Central Research Laboratory. I had the pleasure to hear and question Zurek in a small group at SMU. Dr. Cy Cantrell reminded me a density matrix was a Hermitian operator and therefore must have orthogonal eigenstates. He pointed me to very interesting literature on coherence and Pi pulses. Dr. Wolfgang Rindler, through much discussion, helped me to get my ideas somewhat clearer.

VI. REFERENCES :

-
- [1] A. Albrecht. *Phys. Rev. D*, 48:3768, 1993.
 - [2] M.O. Scully and S.F. Jacobs. *Applied Optics*, 9:2414, 1970.
 - [3] Berthold-Georg Englert, Julian Schwinger, and M.O. Scully. *Foundations of Physics*, 18:1045, 1988.
 - [4] W.H. Zurek(1983). "Information Transfer in Quantum Measurements:Irreversibility and Amplification", p.87. In P. Meystre and M.O. Scully, eds., *Quantum Optics, Experimental Gravitation, and Measurement Theory*. Plenum, New York.
 - [5] W.H. Zurek, S. Habib, and J.P. Paz. *Phys. Rev. Lett.*, 70:1187, 1993.